

АВТОМАТИЗИРОВАННАЯ СИСТЕМА РАСПОЗНАВАНИЯ И КЛАССИФИКАЦИИ ЛИЦ НА ОСНОВЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Рассматривается реализация системы распознавания и классификации лиц на основе компьютерного зрения. Предлагается использование Python, фреймворка FastAPI, библиотеки OpenCV и архитектуры сверточных нейронных сетей для извлечения лицевых признаков. Описана клиент-серверная архитектура модуля, интегрирующая реляционные и нереляционные базы данных (MySQL, Redis) для эффективного хранения и сравнения многомерных векторов.

Введение

В условиях ежедневного роста объемов медианных данных задача автоматической идентификации лиц становится крайне актуальной. Контроль доступа к помещениям с конфиденциальной информацией требует высокотехнологичных решений, так как ручной просмотр видеоданных неэффективен, сопряжен с рисками и высокими трудозатратами. Цель разработанного модуля – полная автоматизация классификации лиц с применением алгоритмов глубокого обучения, способных обрабатывать массивы изображений и потоковые данные в реальном времени.

I. Методы распознавания

Процесс распознавания лиц сводится к преобразованию сегментов изображений в многомерные числовые векторы (эмбединги) с помощью операции свертки с последующей передачей на полносвязный слой. Одной из передовых систем для векторного представления является FaceNet, преобразующая лица в компактные евклидовы пространства [1].

Для определения сходства применяется косинусная метрика, которая, в отличие от евклидова расстояния, устойчива к масштабированию и сохраняет эффективность в высокоразмерных пространствах. Вычисление косинусного расстояния *similarity* производится по формуле:

$$similarity(a, b) = \frac{a \cdot b}{|a| \cdot |b|}$$

где $|a|$ и $|b|$ – нормы векторов a и b . Лицевыми детекторами выступают RetinaFace для статичных кадров и YuNet для потокового видео [2].

II. Архитектура и алгоритмы

Архитектура базируется на модели с использованием «тонкого» клиента (рис. 1 1), что позволяет перенести основные вычисления на сервер и повысить безопасность [3].

Такой подход не только минимизирует требования к аппаратному обеспечению на стороне конечного пользователя, но и существенно упрощает процесс масштабирования системы. Централизованное управление бизнес-логикой на сервере гарантирует оперативную доставку обновлений всем клиентам одновременно, исключая рассинхронизацию версий ПО. Кроме того, хранение биометрических шаблонов и обработка видеопотоков в защищенном серверном контуре значительно снижают риски компрометации

конфиденциальных данных при утере или взломе клиентского устройства.

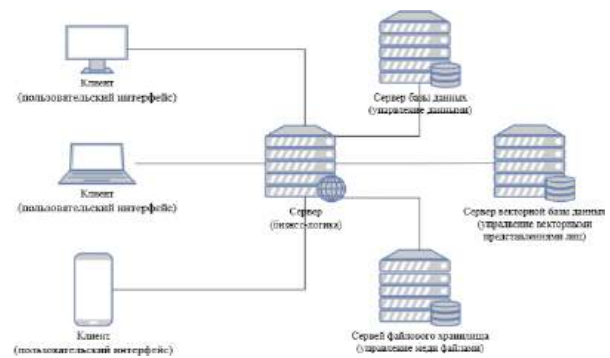


Рис. 1 – Архитектура модуля

Система включает: серверную часть на Python (FastAPI) для асинхронной обработки запросов; клиентское SPA-приложение (React, TypeScript); а также комбинированное хранилище данных – реляционную БД MySQL, in-memory БД Redis для быстрого поиска векторов и FTP-сервер Alpine.

Алгоритм поиска схожих лиц представляет собой последовательность из нескольких ключевых этапов. Первоначально система выполняет строгую валидацию загружаемых файлов и их базовую предобработку с нормализацией. Далее происходит анализ кадра и обнаружение всех присутствующих лицевых сегментов. При успешном обнаружении нейросеть извлекает многомерные эмбединги, которые направляются в Redis для высокоскоростного поиска ближайших соседей. На заключительном этапе осуществляется строгая фильтрация полученных результатов по заданному критерию точности для минимизации ложных срабатываний. Детальная блок-схема описанного процесса представлена на рис. 2.

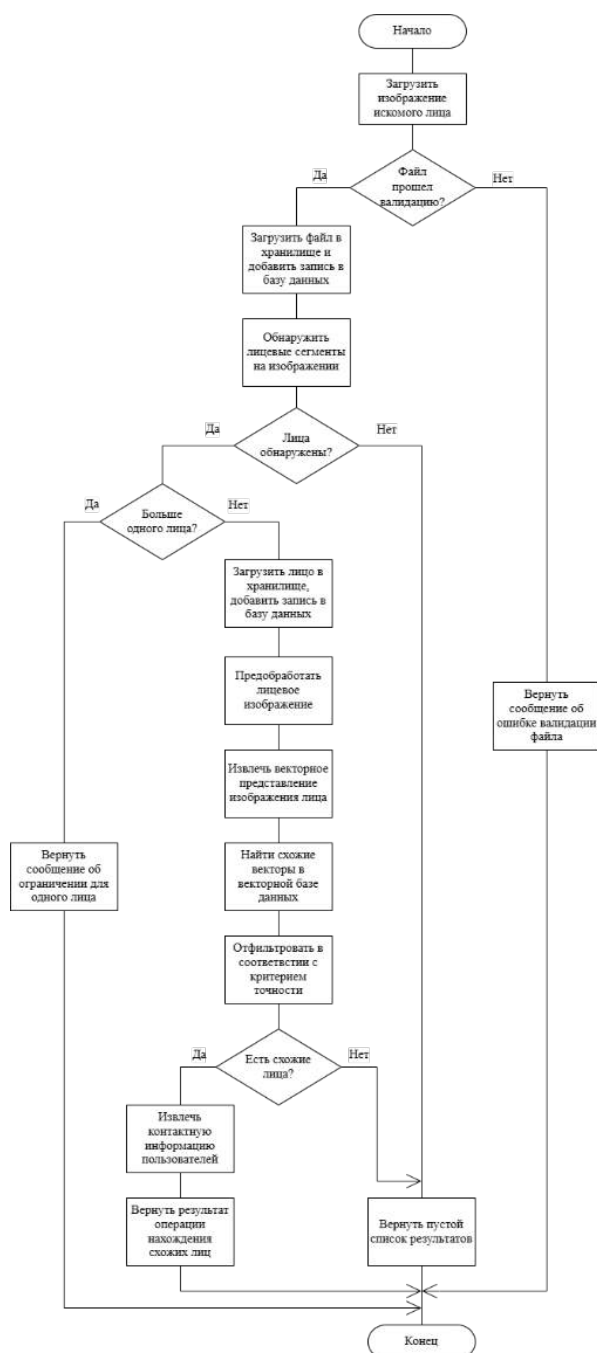


Рис. 2 – Схема поиска схожих лиц

Хаджинова Ксения Александровна, студентка 3 курса факультета информационных технологий и управления, БГУИР, kseniyakhadzhyanova@gmail.com.

Земляник Руслана Владимировна, студентка 3 курса факультета информационных технологий и управления, БГУИР, rzemlanik@gmail.com.

Зубок Роман Александрович, выпускник факультета информационных технологий и управления, БГУИР, renzor.wrk@gmail.com.

Научный руководитель: Хаджинова Наталья Владимировна, старший преподаватель кафедры информационных технологий автоматизированных систем, БГУИР, khajynova@bsuir.by.

Фоновый планировщик CRON выполняет регулярные задачи: синхронизацию FTP с базой данных, очистку неактуальных векторов и покадровую обработку видеопотока с IP-камер.

III. Экономическая эффективность

Внедрение модуля снизило время на ручную идентификацию с 2,5 до 0,4 человеко-часа на одну задачу. Экономия на фонде оплаты труда обеспечила прирост чистой прибыли в размере 32731,86 руб. в год. Оценка рентабельности инвестиций (*ROI*) проводилась по формуле:

$$ROI = \frac{NP}{IC} \cdot 100\%$$

где *NP* – чистая прибыль проекта, *IC* – инвестиционные затраты.

Расчетная рентабельность (*ROI*) составила 52%, а чистый дисконтированный доход за четыре года эксплуатации оценивается в 42228,71 руб. Срок окупаемости равен трем годам, что доказывает высокую экономическую целесообразность проекта.

IV. Выводы

Модуль успешно автоматизирует сбор, обнаружение и классификацию лиц. Использование стека Python, FastAPI, React и нейросетевых детекторов позволяет точно обрабатывать статику и потоковое видео в реальном времени, сводя к минимуму ручной труд. В перспективе планируется интеграция модуля с системами анализа поведения для оценки эмоционального фона сотрудников.

Список литературы

1. Schroff, F. FaceNet: A Unified Embedding for Face Recognition and Clustering / F. Schroff, D. Kalenichenko, J. Philbin // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 815-823.
2. Deng, J. RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild / J. Deng [et al.] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2020. – P. 5203-5212.
3. Лиманова, Н. И. Анализ эффективности современной клиент-серверной архитектуры / Н. И. Лиманова, И. А. Селезнев // Бюллетень науки и практики. – 2022. – Т. 8. – №7. – С. 392-398.